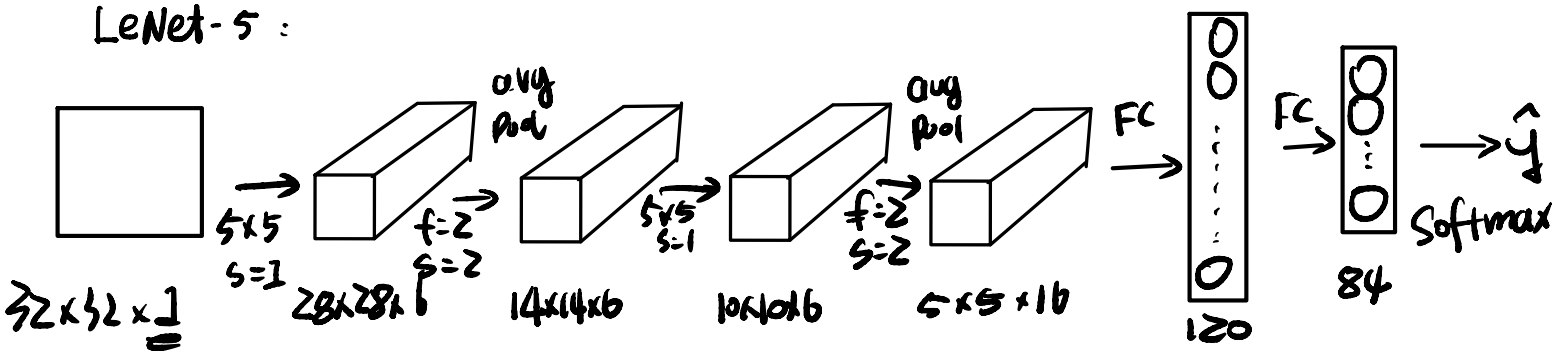
Classic Networks

- LeNet-5
- AlexNet
- VGG

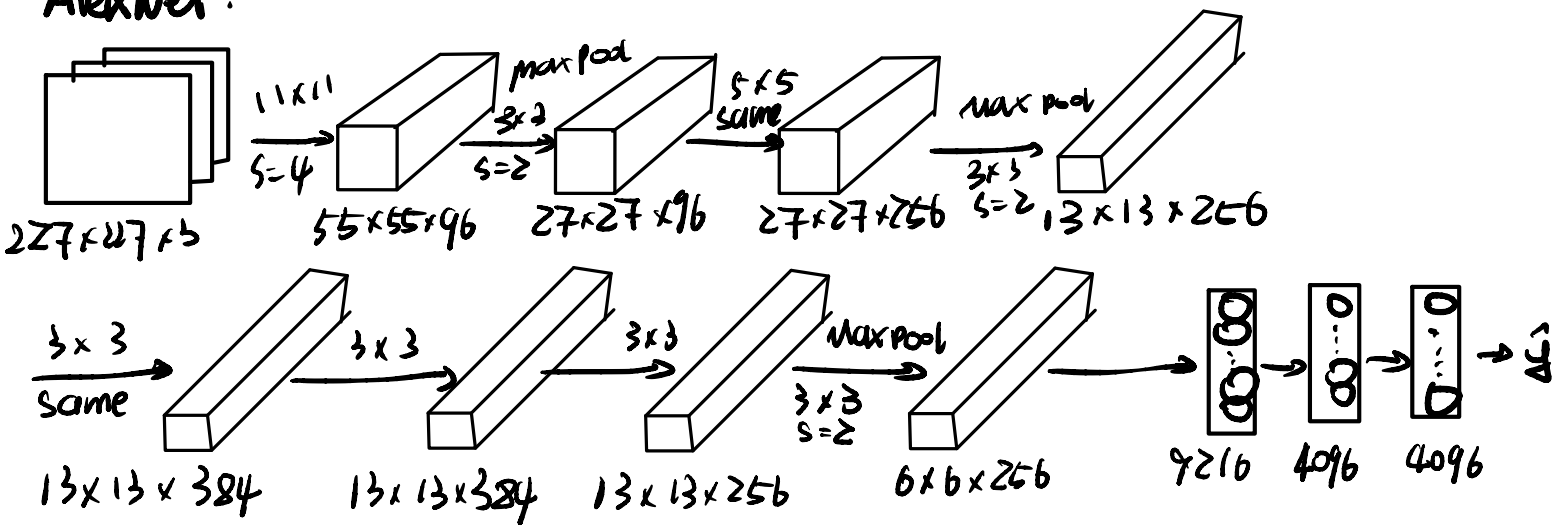
LeNet-5 :



Grey scale image

- Originally designed to perform grey scale digit recognition
- ~ 60k parameters
- $n_H, n_W \downarrow$; $n_C \uparrow$
- Conv → pool → conv → pool → fc → fc → output.

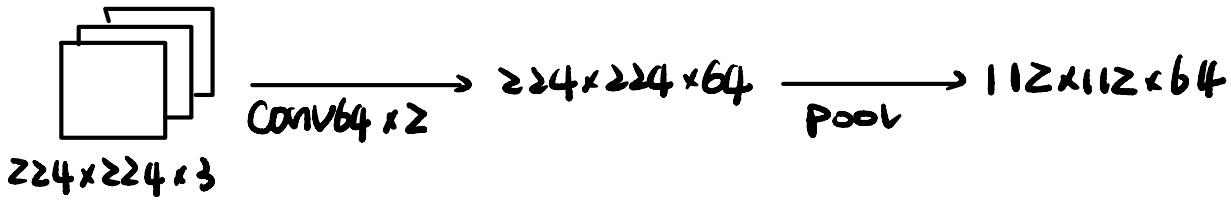
AlexNet :



- Similar to LeNet-5, but much bigger, 60M parameters.
- ReLU
- Multiple GPUs
- Local Response Normalization

VGG-16

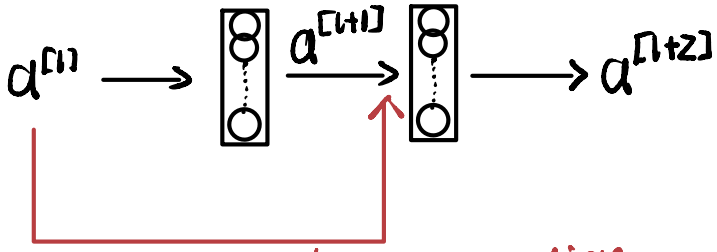
CONV = 3×3 filters, S=1, Same: MaxPool = 2×2, S=2



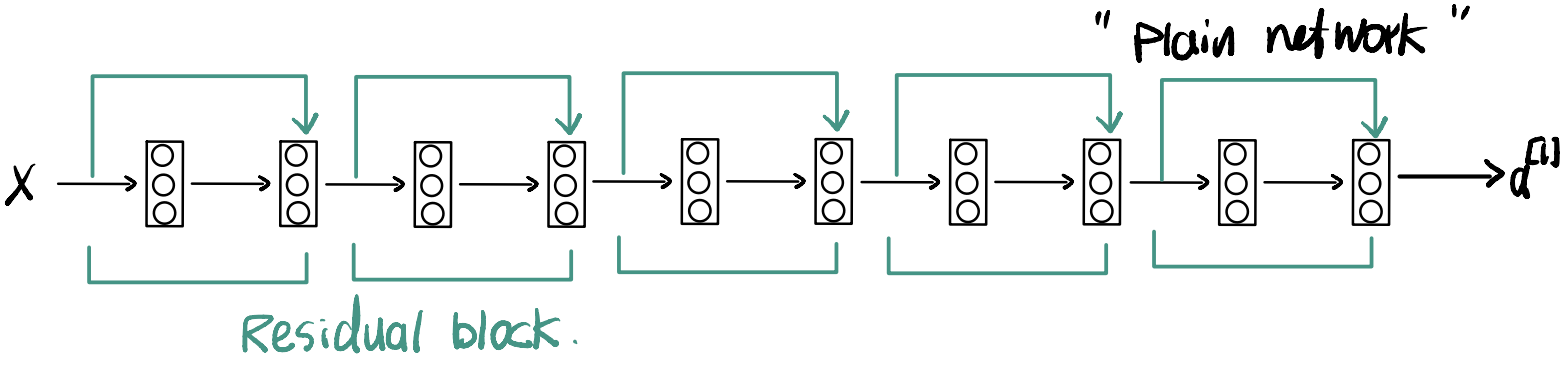
~ 138 M parameters.

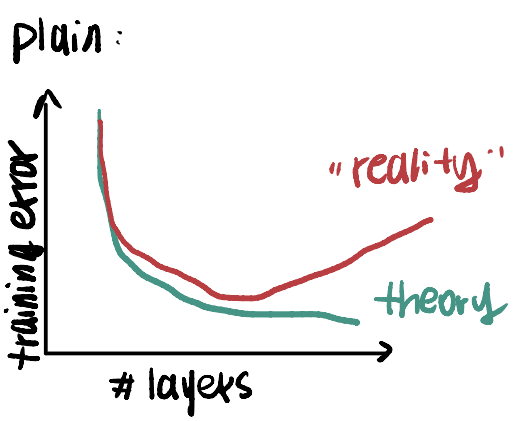
$n_H, n_W \downarrow$; $n_C \uparrow$ systematically.

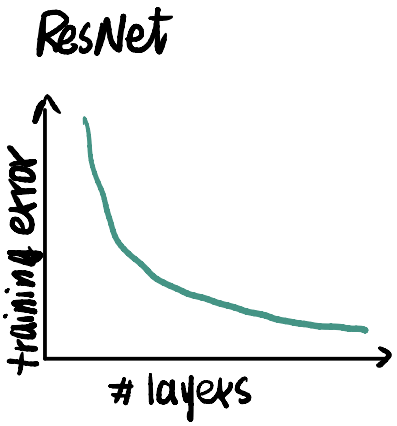
ResNets Residual Networks.



$$a^{[l+2]} = g(z^{[l+2]} + a^{[l]})$$







Network in Network (1×1 convolution)

Example:

$6 \times 6 \times 32 * 1 \times 1 \times 32 = 6 \times 6 \times \#filters$.

for each of the 36 (6×6) input vectors $x \in \mathbb{R}^{32}$, we are applying a weight $w \in \mathbb{R}^{32}$ to it. this is the same effect of a neuron in the neural network. if we have multiple filters, this can become a fully connected network. Thus 1×1 convolution is also called network in networks.

other effects:

Example: $28 \times 28 \times 192 \xrightarrow[\substack{\text{conv} 1 \times 1 \\ 32}]{\text{ReLU}} 28 \times 28 \times 32$

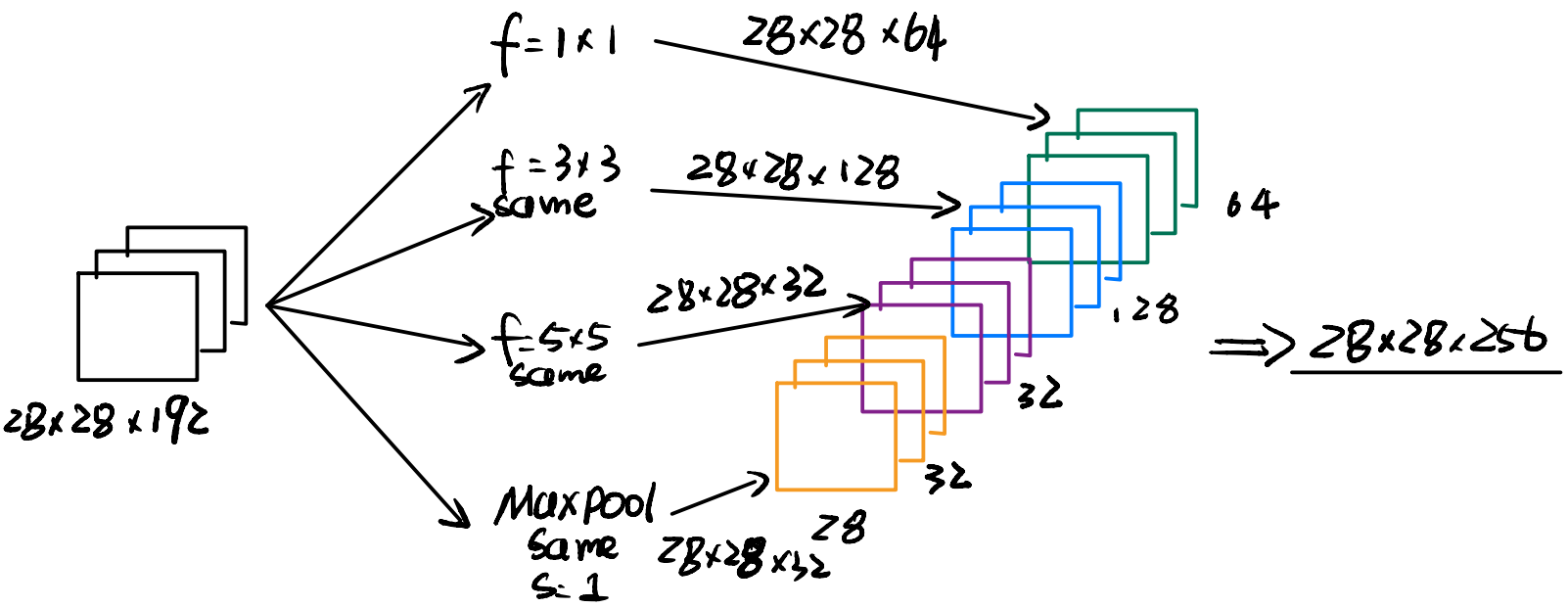
1×1 conv layer can be used to shrink depth. (# channel).

$28 \times 28 \times 192 \xrightarrow[\substack{\text{conv} 1 \times 1 \\ 192}]{\text{ReLU}} 28 \times 28 \times 192$

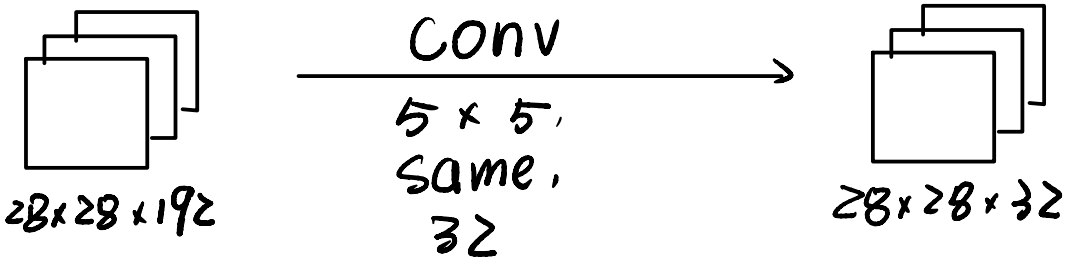
it can also be used to just add a layer of nonlinearity.

Inception Network

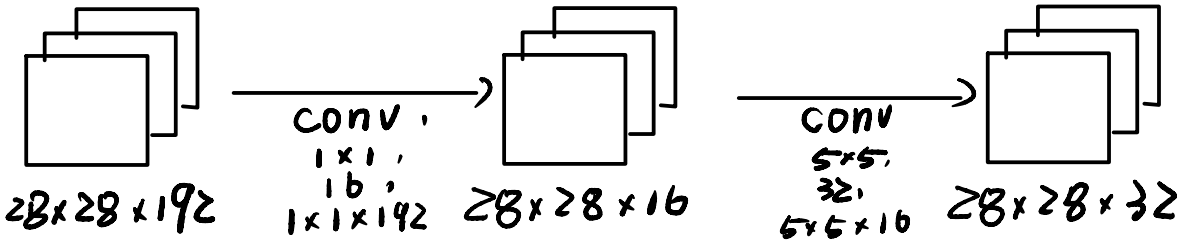
motivation : use convolutions of all sizes.



problem of computational cost

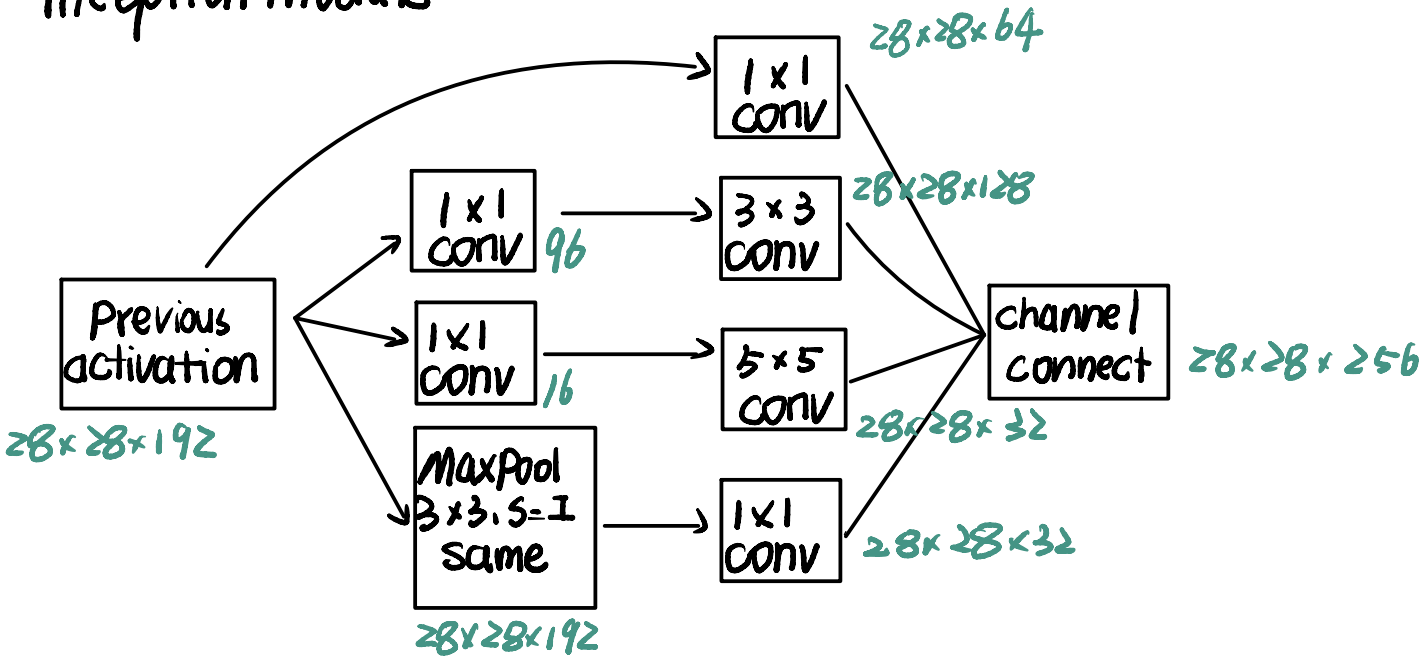


Computations = $28 \times 28 \times 32 \times 5 \times 5 \times 192 = 120M$



Computations = $\underbrace{28 \times 28 \times 16 \times 192}_{2.4M} + \underbrace{28 \times 28 \times 32 \times 5 \times 5 \times 16}_{10M} = 12.4M$

Inception module

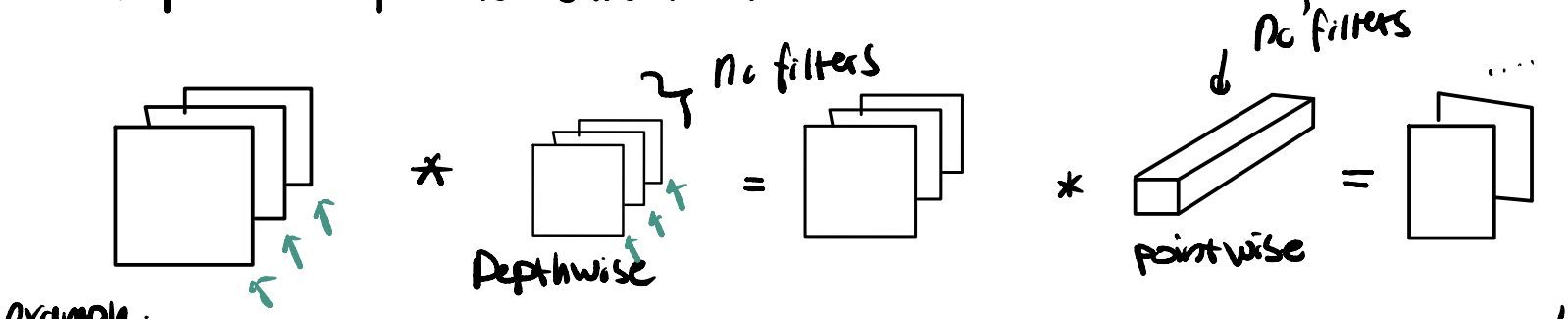


MobileNet : designed for low computing environments, e.g. mobilephone

motivation :

Normal Convolution :
computational cost = # filter params × # filter positions × # of filters

Depthwise Separable Convolution :

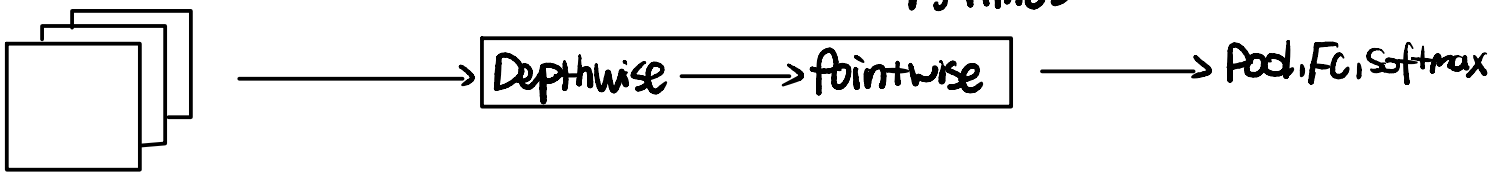


Cost Summary :

$$\frac{\text{Cost of normal convolution}}{\text{\# of depthwise seperable conv}} = \frac{1}{n_c'} + \frac{1}{f^2}$$

~ 10 times cheaper

MobileNet v1



MobileNet V2